

# 인간의 시각 특성을 고려한 Sub-Band 부호화에서 MTF 비교에 관한 연구

## (On the Comparison of MTF in Sub-Band Coding Technique Employing the Human Visual System)

金 容 寬\*, 朴 涉 亨\*, 李 商 郁\*

(Yong Kwan Kim, Seop Hyeong Park, and Sang Uk Lee)

### 要 約

본 논문에서는 sub-band 부호화 기법에 인간 시각 시스템 모델로 이미 제안된 대표적인 4 가지의 MTF(modulation transfer function) 즉, Sakrison, Nill, Ngan, Rao의 MTF들을 적용하여 그 결과들을 주관적인 기준에서 비교, 평가하였다.

Sub-band 부호화 과정에서 영상 신호의 주파수 대역을 똑같은 크기의 16개 sub-band들로 분할하였다. 이와 함께 저주파 대역을 세분하는 28 sub-band 분할 형태를 위하여 4가지 MTF들의 성능 비교가 용이하도록 하였다. 각 sub-band에 대한 부호화 과정에서, 인간 시각 시스템의 MTF로부터 각 sub-band의 가중치를 구하여 가중 평균 자승 오차를 최소화하도록 비트를 할당하였다. 제일 낮은 공간 주파수 성분을 갖는 sub-band에는 DPCM을, 나머지 sub-band들에는 PCM을 이용하여 부호화하였다. 재생된 영상들에 대한 주관적인 평가로부터, Rao의 MTF를 반영한 부호화 결과가 가장 우수한 성능을 나타내었으며, 다음으로 Ngan, Nill, Sakrison, 그리고 인간 시각 시스템의 MTF를 고려하지 않은 경우의 순으로 나타났다.

### Abstract

In this paper, attempts have been made to compare the subjective performance of sub-band coding (SBC) techniques incorporating four representative modulation transfer functions (MTF's) of the human visual system(HVS): Sakrison, Nill, Ngan, and Rao.

In SBC, the frequency band of input signal is split into 16 equal sub-bands. In addition, 28 sub-band splitting scheme which splits more sub-bands in low spatial frequency is considered to compare the 4 MTF's effectively. In encoding of each sub-band, the weight of each sub-band obtained from the MTF of HVS is applied to bit allocation process which minimize the weighted mean square error (WMSE). The differential pulse code modulation (DPCM) coder is used to encode the lowest sub-band and the pulse code modulation (PCM) coder is used for the rest of sub-bands. It is found that the images incorporating the MTF of Rao yields the best results in subjective criteria, followed by Ngan, Nill, Sakrison, and the images not employing the HVS.

\*正會員, 서울대학교 制御計測工學科 信號處理研究室  
(Dept. of Control and Instrumentation Eng., Seoul  
Nat'l Univ.)

接受日字: 1990年 2月 14日

### I. 서 론

영상 부호화의 목적은 원 영상을 충실히 재생하  
면서 그에 필요한 데이터량을 가능한 한 줄이는데 있

다. 1960년대 초부터 정보 이론에 기초하여 발전되어 온 영상부호화 기법의 데이터 감축비가 1980년대 초에 이르러 10:1 정도의 한계점에 도달하게 되었다.<sup>[1]</sup> 영상 부호화의 최종적인 결과는 일반적으로 인간의 시각에 의해서 판단되므로, 영상 부호화에서 현재까지 알려진 인간 시각 시스템(human·visual system: HVS)을 고려하여 영상 데이터의 감축비를 높이려는 시도가 1980년대 이후부터 활발히 진행되고 있다.<sup>[1]</sup>

인간의 시각 기관은 영상을 공간 주파수 성분에 따라 각기 독립적으로 처리하는 구조를 가지고 있으며 각각의 공간 주파수 성분에 대한 시각 감도가 다른 것으로 알려져 있다.<sup>[2]</sup> 일반적으로 인간 시각 시스템의 공간 주파수에 따른 감도 변화 특성을 주파수 영역에서 정의되는 MTF(modulation transfer function)로 모델링한다. 영상 부호화에서 인간 시각 시스템의 MTF를 적용하는 목적은 인간의 시각이 민감한 감도는 갖는 공간 주파수 대역 신호는 충실하게 부호화하고 그렇지 않은 부분에서는 영상의 표현에 필요한 정보량을 제한함으로써 한정된 정보량으로 인간의 시각에 충실한 부호화를 행하여 주기 위한 것이다. 이러한 연구는 변환 부호화를 중심으로 이루어져 왔다.<sup>[3-6]</sup> Manno와 Sakrison은<sup>[3]</sup> 영상 부호화에 인간 시각 시스템 모델을 최초로 도입하였다. 그들은 전체 신호를 DFT(discrete fourier transform) 변환하여 얻은 변환 계수들에 인간 시각 시스템의 MTF를 곱하고, 일정한 임계값을 넘는 계수들을 이용하여 신호를 재생하는 방법을 통하여 인간 시각 시스템의 MTF를 구하였다. 그 이후, Nill은<sup>[4]</sup> 영상 부호화와 영상의 화질 평가 방법에서 인간 시각 시스템의 MTF를 cosine 변환 영역에 적용할 수 있도록 하는 함수를 제안하였다. Ngan 등<sup>[5]</sup>은 Nill이 제안한 함수를 MTF에 곱하여 얻은 가중치들을 DCT(discrete cosine transform)변환 계수들에 가중시킴으로써 전송율 0.2-0.3bpp에서 만족할 만한 재생 영상을 얻었다고 발표하였다. 최근에 Rao는 DCT 변환 부호화에 인간 시각 시스템의 MTF를 결합한 영상의 점진적 전송 방법에 관한 연구를 하여,<sup>[6]</sup> 낮은 전송율에서 주관적인 기준에서 우수한 결과를 얻었다고 주장하였다. 그러나 변환 부호화는 영상 신호를 공간상에서 부블럭 단위로 변환하여 부호화하므로 낮은 전송율에서는 블럭 현상(blocking effect)이 발생하는 단점이 있다. 반면에, sub-band 부호화 기법은 신호를 필터 뱅크(filter bank)를 통하여 공간 주파수 대역별로 분할하여 부호화하므로 변환 부호화에서 발생할 수 있는 블럭 현상이 나타나지 않는 장점이 있으며,<sup>[5]</sup> 공간 주파수에 따른 인간의 시각 특성을 효과

적으로 반영할 수 있는 부호화 방법이라 할 수 있다. Perkins 등은 sub-band 부호화 기법에서 인간 시각의 공간 주파수에 따른 시각 감도 특성을 비트 할당 과정에 도입함으로써 주관적으로 우수한 평가를 얻을 수 있는 부호화 방법을 제안하고 있다.<sup>[8]</sup> 그들은 [2]에서 제안하는 인간 시각의 공간 주파수 감도 특성을 2차 함수로 근사화시켜 부호화에 적용하였다. 일반적으로 인간 시각 시스템의 MTF가 어떠한 공간 주파수에서 최대값을 갖는가에 따라 MTF를 고려한 출력 영상의 차이가 심하게 나타난다.<sup>[3]</sup> 따라서 이미 제안된 여러가지의 인간 시각 시스템의 MTF들로부터 sub-band 부호화에 가장 적합한 MTF를 선택할 필요가 있다.

본 논문에서는 이미 제안된 4가지의 MTF 들로부터 구한 각 sub-band의 가중치를 sub-band 신호에 곱하여 가중 평균 자승 오차(weighted mean square error)를 최소화 하도록 비트를 할당하였다. 또한 부호화 결과들을 주관적으로 비교, 평가하였으며 그로부터 sub-band 부호화 기법에 가장 적합한 인간 시각 시스템의 MTF를 선정하였다.

## II. 인간 시각 시스템의 MTF

인간 시각 특성에 관한 연구 결과들에서는 인간의 시각이 신호의 공간 주파수에 따라서 감도가 변하는 특성을 대조 감도 함수(contrast sensitivity function)로 나타낸다.<sup>[2,3,9]</sup> 인간 시각의 광학적인 불완전성에 의한 저주파 대역 통과 필터링(low-pass filter) 효과와 시각 신경에서의 측방 억제(lateral inhibition) 현상에 의한 고주파 대역 통과 필터(high-pass filter) 역할등에 의하여, 대조 감도 함수는 약 3-4.5(cycles/degree)에서 최대값을 나타내는 중간 대역 통과 필터(band-pass filter)의 특성을 갖는 것으로 알려져 있다.<sup>[2,9,10]</sup> Manno와 Sakrison은 시각 특성에 의한 대조 감도 함수의 최대값이 1이 되도록 정규화하여 식(1)과 같은 인간 시각 시스템의 MTF를 제안하였다.<sup>[3]</sup> 그후에 계속해서 변환 부호화 기법에서 Nill, Ngan, Rao 등은 각각 식(2), 식(3), 식(4)와 같은 MTF 들을 제안하였다.<sup>[4-6]</sup>

$$H(f) = 2.6(0.0192 + 0.114f) \exp(-0.114f^{1.1}),$$

$$f_0 = 8(\text{cycles/degree}) \quad (1)$$

$$H(f) = (0.2 + 0.45f) \exp(-0.18f), \quad f_0 = 5.1(\text{cycles/degree})$$

$$(2)$$

$$H(f) = (0.31 + 0.69f) \exp(-0.29f), \quad f_0 = 3(\text{cycles/degree})$$

$$(3)$$

$$H(f) = 2.46(0.1 + 0.25f) \exp(-0.25f), \quad f_0 = 3.75(\text{cycles/degree})$$

$$(4)$$

위의 식들에서  $f_0$ 는 최대의 감도를 갖는 공간 주파수를 나타낸다. 그림 1은 4가지 MTF들을 비교하여 나타내고 있다. 본 논문에서는 위의 4가지 MTF들을 sub-band 부호화에 적용하여 그 특성을 비교, 분석하였다.

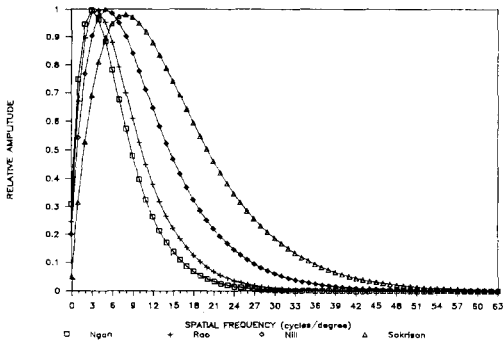


그림 1. 4 가지 MTF 함수들의 비교  
Fig. 1. A comparison of 4 MTF's.

III. 시각 특성에 근거한 Sub-Band 부호화

Sub-band 부호화의 기본적인 개념은 신호를 주파수 대역별로 분할하여 각각의 주파수 대역을 독립적으로 부호화하는 것으로 다음과 같은 장점이 있다. 첫째, 각 주파수 대역의 특성에 적합한 부호화 방법으로 데이터의 압축 효율을 높일 수 있다. 둘째, sub-band 신호를 부호화할 때 발생하는 오차가 해당 주파수 대역에 국한되어 나타나므로 복원후 전체 신호에 미치는 영향이 제한된다. 셋째, 분할된 각각의 sub-band에 할당되는 비트수를 조절함으로써 양자화 잡음 스펙트럼(noise spectrum)의 형태를 변화시켜 주관적으로 우수한 신호를 재생시킬 수 있다.

Sub-band 부호화는 1976년 Crochiere 등에 의해 처음으로 소개<sup>[11]</sup>된 이후 음성 신호의 부호화에 널리 이용되어 왔다. 1984년 M. Vetterli는 다차원 신호를 sub-band들로 분할할 수 있는 방법을 제시하였다.<sup>[12]</sup> 이어서 1986년 Woods와 O'Neil에 의하여 영상에서의 sub-band 부호화가 연구되기 시작하였으며<sup>[13]</sup> 현재까지 sub-band 부호화에 관한 연구들이 많이 이루어져 왔다.<sup>[14-16]</sup>

다음에서는 본 논문에서 사용한 주파수 대역 분할 방법과 각 sub-band 신호에 대한 부호화 기법 및 인간 시각 시스템의 MTF를 고려한 비트 할당 방법에 대하여 설명한다.

1. 주파수 대역 분할 및 부호화 방법

Sub-band 부호화에서 가장 먼저 고려해야 할 것은 신호의 주파수 대역을 분할하는 것으로서, 일반적으로 QMF를 사용하여 주파수 대역을 분할, 합성한다.<sup>[17]</sup> 그림2는 2차원 영상 신호를 4개의 주파수 대역으로 분할 및 합성하는 과정을 나타내고 있다. 여기서는 1차원 QMF를 이용하여 영상 신호를 행 방향으로 2개의 대역으로 분할한 뒤에 간축 과정을 거치고 다시 열 방향으로 동일한 과정을 거치는 4개의 주파수 대역 분할 과정과 그 역과정인 합성 과정을 통하여 신호를 재생한다.<sup>[12]</sup> 이와 같은 4개의 sub-band 분할 과정을 분할한 주파수 대역에 대하여 반복하는 트리 구조(tree structure)를 통하여 공간 주파수 대역을 세분할 수 있다. 그러나 QMF 필터링 과정에서 신호의 가장자리(boundary)에 왜곡이 발생하며, 이러한 왜곡을 완화시키기 위하여 QMF 필터링 과정의 모든 입력 신호들을 적절한 방법으로 확장시킬 필요가 있다.<sup>[17]</sup> 본 논문에서는 순환 확장(circular extension)과 대칭 확장(symmetric extension)방법, 그리고 주변 확장(replication of edge value)방법을 비교하였다. 0.5 bpp의 전송율(bit rate)로 LENA 영상을 부호화한 결과, PSNR 값이 각각 31.91 dB, 32.18 dB, 32.24dB로 나타났으며 주관적인 입장에서 관찰하였을 때도 [7]에서 제안하는 방법과 같이 주변 확장 방법이 가장 우수한 것을 확인하였다.

본 논문에서는 SNR 기준에서 우수한 성능을 갖는 것으로 발표된<sup>[15]</sup> 그림 3(a)에서와 같이 16개의 동일한 대역 폭을 갖는 sub-band 분할 형태를 취하

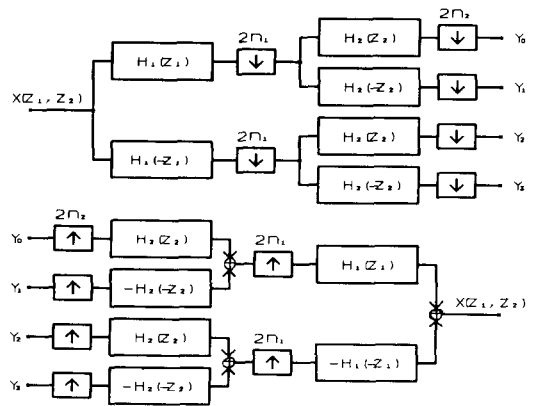


그림 2. 4 채널 separable QMF 필터 뱅크의 분할/합성 과정

Fig. 2. 4 sub-band analysis/synthesis using separable QMF bank.

었다. 이와함께 낮은 주파수 대역에서 많은 변화를 나타내는 4가지 MTF들간의 비교가 타당하게 이루어질 수 있도록 낮은 공간 주파수 대역을 세분하는 28 sub-band 분할 형태를 이용한 부호화를 아울러 고려하였으며 그 영향에 대하여 고찰하였다. 28 sub-band 분할 형태와 같이 공간 주파수 대역을 세분하는 것은 분할에 따르는 왜곡을 무시한다면, 부호화 과정의 비트 할당 과정에서 더욱 유리한 조건이 될 수 있다.<sup>[16]</sup> 한편 인간의 시각은 공간 주파수  $f$ 와 그로부터 3배의 간격을 갖는 공간 주파수  $3f$ 를 서로 독립적으로 처리하는 것으로 알려져 있다.<sup>[2,22]</sup> 따라서 저주파 대역을 세분하는 방법에 의해 인간 시각 시스템 모델에 더욱더 근접할 수 있을 것으로 생각된다.

신호의 주파수 대역을 분할한 뒤에 각 sub-band 신호의 특성에 맞는 부호화 기법을 사용하는 것이 바람직하다.<sup>[14,15]</sup> sub-band 신호들의 자기 상관 계수를 조사한 결과에서 제일 낮은 주파수 대역 신호의 상관도가 높게 나타나는 반면에, 나머지 높은 주파수 대역 신호들의 상관도는 낮은 값을 나타내므로 본 논문에서는 각각 DPCM과 PCM 부호화기를 사용하였다. DPCM 부호화 과정에서는 3개의 주위 화소들을 이용하는 다음과 같은 선형 예측기를 사용하였다.<sup>[14]</sup>

$$x(m, n) = A \cdot x(m-1, n) + B \cdot x(m, n-1) + C \cdot x(m+1, n-1) \quad (5)$$

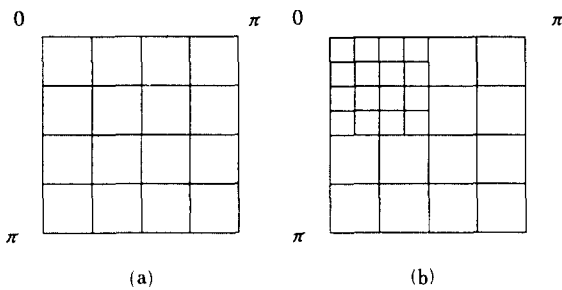


그림 3. 512×512 LENA 영상에 대한 sub-band 분할 형태 및 합성결과(PSNR)  
 (a) 16 sub-band 분할 형태(합성 결과: 50.80dB)  
 (b) 28 sub-band 분할 형태(합성 결과: 47.39dB)

Fig. 3. Sub-band splitting schemes and the PSNR's of synthesized LENA images.  
 (a) 16 sub-band splitting scheme (50.80dB)  
 (b) 28 sub-band splitting scheme (47.39dB).

여기서  $A$ 는 0.5,  $B$ 는 0.25,  $C$ 는 0.25 이며,  $m$ 과  $n$ 은 각각  $X, Y$  방향으로의 화소 위치를 나타낸다.

2. MTF를 고려한 비트 할당 방법

Sub-band 부호화 과정에서  $i$ 번째 sub-band 신호를 스칼라 양자화할 때, 해당 sub-band 신호의 평균 자승 오차(mean square error)는 다음의 식과 같이 나타낼 수 있다.<sup>[17]</sup>

$$D_i = k_i \cdot 2^{-2n_i} \cdot \sigma_{x_i}^2 \quad (6)$$

여기서  $D_i$ ,  $\sigma_{x_i}^2$ ,  $n_i$ ,  $k_i$ 는 각각 sub-band  $i$ 의 양자화에 의한 평균 자승 오차, 신호의 분산, 할당된 비트수, 그리고 양자화 상수를 나타낸다. Sub-band 부호화의 비트 할당 과정에서 문제가 되는 것은 제한된 정보량으로 양자화에 의한 전체 재생 영상의 오차를 최소로 하도록 하는 것이다. 즉, 부호화한 영상 신호의 왜곡 측정자가 평균 자승 오차이고, 신호를  $N$ 개의 sub-band들로 분할하는 경우에, 다음의 식(7)과 식(8)로 나타낼 수 있다.<sup>[17]</sup>

$$\text{Minimize } D = \sum_{i=1}^N D_i \quad (7)$$

$$\text{subject to } B = \frac{1}{T} \sum_{i=1}^N T_i \cdot B_i \text{ (bits/pixel)} \quad (8)$$

여기서,  $D$ 와  $B$ 와  $T$ 는 각각 전체 재생 영상에 대한 평균 자승 오차와 평균 전송율, 전체 신호의 주파수 대역 폭을 각각 나타내며, 따라서  $D_i$ 와  $B_i$ 와  $T_i$ 는  $i$ 번째 sub-band의 평균 자승 오차와 평균 전송율, 그리고 주파수 대역폭을 나타낸다.

한편 인간 시각 시스템을 근사적으로 선형 시스템으로 가정하면,<sup>[4]</sup> MTF를 시각 시스템의 전달 함수의 크기 특성으로 간주할 수 있다.<sup>[18]</sup> 본 논문에서는 sub-band의 중앙 주파수 영역에서 MTF 함수값을 가중치로서 해당 sub-band 신호에 직접 곱하였다. 따라서 가중 평균 자승 오차(weighted mean square error)를 최소화하는 최적 비트 할당 식을 다음과 같은 식(8)과 식(9)로 표현할 수 있다.

$$\begin{aligned} \text{Minimize } D &= \sum_{i=1}^N D_i \cdot W_i^2, \quad (9) \\ &= \sum_{i=1}^N k_i \cdot 2^{-2n_i} \cdot (W_i \cdot \sigma_{x_i}^2)^2 \end{aligned}$$

subject to (8).

식(9)에서 각 sub-band의 가중치  $W_i$ 를 구하기 위해서는 먼저 영상에 대한 관찰 조건이 정해져야 한다. 영상이  $M \times M$ 의 화소로 이루어지고, 영상에 대한 수

평 및 수직 관찰 각도를  $\alpha$ 로 나타낼 때, 각 방향으로 최대로 나타날 수 있는 공간주파수  $f_{max}$ 는 다음의 식(10)과 같다.

$$f_{max} = (M/2)/\alpha \text{ (cycles/degree)} \quad (10)$$

영상 신호의 주파수 대역을 수평 또는 수직 방향으로 동일한 대역 폭을 갖는 K개로 분할하는 경우, 각 방향으로의 sub-band 대역 폭은  $f_{max}/K$ 가 된다. 각 sub-band의 중앙 공간 주파수에서의 X, Y 방향으로의 공간 주파수를  $f_x, f_y$ 라고 할 때, 해당 sub-band의 중앙 공간 주파수는 다음의 식(11)과 같이 나타낼 수 있다.

$$f_r = \sqrt{f_x^2 + f_y^2} \quad (11)$$

본 논문에서는 각 sub-band의 중앙 공간 주파수  $f_r$ 에서 등방성으로 가정한 MTF로 부터 함수값을 취하고 그것들의 최대값을 1로 정규화하여 해당 sub-band의 가중치로 정하였다. 예를 들면 표1은 Rao의 MTF로부터 구한 16 sub-band 및 28 sub-band 들에 해당하는 가중치들을 나타내고 있다. 식(8)과 식(9)와 같이 표현되는 최적 비트 할당 문제를 해결하는 방법으로, Lagrangian multiplier를 이용하는 방법<sup>[17]</sup>과 marginal analysis 비트 할당 방법<sup>[19]</sup>이 가장 널리 사용된다.

본 논문에서는 일반적인 marginal analysis 비트 할당 방법을 수정하여 적용하였다. 실제의 부호화 과정에서 높은 주파수 대역에 낮은 비트(혹은 레벨)가 할당되므로 짝수 레벨 양자기를 사용하여 영상을 합성하는 경우, 재생 영상이 눈에 거슬리는 현상을 확인하였고, 또한 양자기에 비트수가 할당되는 경우 양자기의 레벨 간격을 세밀히 조정할 수 없으므로, 4가지 MTF들을 비트 할당 과정에 적용하여 비교하는 것이 제한될 수 있다. 따라서 본 논문에서의 모든 양자기는 홀수 레벨 양자기를 사용하였으며 각 양자기에 정수의 비트수를 할당하는 방법을 수정하여 레벨 할당을 행하였다. 식(6)을 레벨에 의한 식으로 수정하여 다음과 같이 나타낼 수 있다.

$$D_i = k_i \cdot L_i^{-2} \cdot \sigma_{x_i}^2 \quad (12)$$

여기서  $L_i$ 는 sub-band i에 할당된 레벨 수를 나타낸다. 식(8)과 식(9), 그리고 식(12)로부터 레벨로 표현되는 marginal 비트 할당 알고리즘을 다음과 같이 나타낼 수 있다.

먼저 N개의 sub-band들이 있다고 가정하고 모든 sub-band의 양자화 상수  $k_i = k$ 로 가정하였으며 j는 알고리즘의 반복 횟수를 나타낸다.

표 1. Rao의 MTF 가중치  
(a) 16 sub-band 분할 형태  
(b) 28 sub-band 분할 형태

Table 1. Weights of Rao's MTF.

(a) 16 sub-band splitting scheme,  
(b) 28 sub-band splitting scheme.

|      |       |      |      |       |
|------|-------|------|------|-------|
| 0    |       |      |      | $\pi$ |
| 1.00 | 0.49  | 0.16 | 0.04 |       |
| 0.49 | 0.27  | 0.10 | 0.03 |       |
| 0.16 | 0.10  | 0.04 | 0.01 |       |
| 0.04 | 0.03  | 0.01 | 0.00 |       |
|      | $\pi$ |      |      |       |

(a)

|      |       |      |      |      |       |
|------|-------|------|------|------|-------|
| 0    |       |      |      |      | $\pi$ |
| 1.00 | 0.99  | 0.69 | 0.41 | 0.16 | 0.04  |
| 0.99 | 0.83  | 0.58 | 0.35 |      |       |
| 0.69 | 0.58  | 0.41 | 0.26 | 0.10 | 0.03  |
| 0.41 | 0.35  | 0.26 | 0.18 | 0.04 | 0.01  |
| 0.16 |       | 0.10 | 0.04 | 0.04 | 0.01  |
| 0.04 |       | 0.03 | 0.01 | 0.01 |       |
|      | $\pi$ |      |      |      |       |

(b)

[단계 1]  $L_{i,j} = 1$ , 이때  $1 \leq i \leq N, j = 1$ , 그리고  $L_{i,j}$ 는 i번째 sub-band에 j번째 반복 과정에서 할당된 레벨수를 나타낸다.

[단계 2]  $L_{i,j} = L_{i,j-1} + 2$ , 이때  $i = \max^{-1} \Delta_i$ , 여기서  $\Delta_i$ 는 다음의 식 (13)과 같다.

$$\Delta_i = k[(L_{i,j-1})^{-2} - (L_{i,j} + 2)^{-2}] (W_i \sigma_i)^2 \quad (13)$$

위의 식 (13)에서  $\Delta_i$ 는 sub-band i에 2 레벨(level) 추가에 의한 평균 자승 오차의 감소분을 나타낸다.

[단계 3]  $|\sum_{i=1}^N \log_2(L_{i,j}) - \text{total bits}| < t \quad (14)$

[단계 3]에서는 식 (14)의 조건을 만족시키지 못하는 경우,  $j \rightarrow j + 1$ 로 하고, [단계 2]로 되돌아 간다. 식 (14)의 t는 전체 할당 비트수의 임계 오차를 나타내며 실험적으로 0.01로 정하였다.

### III. 시뮬레이션 결과 및 비교 검토

인간의 시각 시스템 모델로 제안된 대표적인 4가지 MTF들을 sub-band 부호화에 적용하고 그 결과를 주관적인 성능 평가 방법에 의해서 비교하였다. 또 sub-band 분할 형태에서 16개의 똑같은 sub-band

분할 방법과 낮은 주파수 대역을 세분하는 28 sub-band 분할 방법에 의한 시뮬레이션 결과들을 비교하였다. 실험 영상은 흑백 정지 영상으로 512×512의 화소수와 화소당 256 등급의 밝기를 갖는 LENA 영상과 BABOON 영상으로 하였다. 본 논문에서는 Johnstone의 16C형 QMF를 사용하여,<sup>[20]</sup> [12]에서 제안한 가분리(separable) 2차원 QMF 필터 뱅크를 구성하였다. 제일 낮은 주파수 대역에는 DPCM 부호화기를, 나머지 대역들에는 PCM 부호화기를 사용하였으며, 각각 Laplacian 및 generalized Gaussian 함수의 계수가 0.5인 화률 밀도 함수를 가정한 Max-Lloyd 양자기를 사용하였다.

영상의 부호화는 0.25 bpp, 0.5 bpp, 0.75 bpp 에서 행하였으며 각 전송률에서 MTF를 적용하여 부호화한 결과들과 적용하지 않은 결과를 함께 비교하였다. 최종적인 재생 영상들은 총 10명에 의해 주관적으로 평가되었다. 영상을 평가할 때, 모든 평가자들은 영상 높이의 6배의 관찰거리<sup>[21]</sup>인 162cm의 거리를 유지하였다. 이때, 식(10)의  $\alpha$ 는 9.53°이고  $f_{max}$ 는 26.87 (cycles/degree)이다. 주위의 밝기는 밝은 상태를 유지하였으며(40W 형광등×4개), 영상은 Gould사의 Model 7141C19 모니터상에서 관찰하였다. 일반적으로, 객관적인 평가 기준인 PSNR은 인간의 주관적

인 영상 평가 기준을 폭넓게 반영하지 못하는 것으로 알려져 있다.<sup>[21]</sup> 표 2는 부호화된 영상의 화질 저하 정도를 7가지로 정하는 주관적인 평가 기준을 나타낸다.<sup>[19]</sup> 실제의 평가에 있어서 각 전송률에서 MTF 들을 고려한 영상들과 고려하지 않은 영상에 대하여 원영상과 비교하며 impairment scale을<sup>[19]</sup> 부과하도록 하였다. 표 3 과 표 4 는 LENA와 BABOON 영상의 주파수 대역을 28 sub-band와 16 sub-band들로 분할하는 경우에 있어서 4가지의 MTF들을 적용한 재생 영상들과 적용하지 않은 재생 영상에 대한 주관적인 평가 결과와 객관적인 평가 결과를 나타내고 있다. 표 3 과 표 4 에서 알 수 있듯이, 몇가지 경우에는

표 2. Impairment scale

Table 2. Impairment scale.

|  |     |
|--|-----|
| Not noticeable                                   | (1) |
| Just noticeable                                  | (2) |
| Definitely noticeable but only slight impairment | (3) |
| Impairment not objectionable                     | (4) |
| Somewhat objectionable                           | (5) |
| Definitely objectionable                         | (6) |
| Extremely objectionable                          | (7) |

표 3. 28 sub-band 분할 형태에 의한 재생 영상의 평가 결과

Table 3. Performance comparison of 28 sub-band splitting scheme.

| CONDITION       |               |           | NO WEIGHT | NGAN  | RAO     | NILL    | SAKRISON |      |
|-----------------|---------------|-----------|-----------|-------|---------|---------|----------|------|
| LENA<br>IMAGE   | 0.25<br>(bpp) | I. S      | MEAN      | 3.50  | 2.40    | (=NGAN) | 4.10     | 5.10 |
|                 |               |           | S. D      | 1.28  | 0.92    | (=NGAN) | 1.04     | 0.94 |
|                 |               | PSNR (dB) | 29.57     | 29.86 | (=NGAN) | 29.23   | 28.81    |      |
|                 | 0.5<br>(bpp)  | I. S      | MEAN      | 4.00  | 2.00    | (=NGAN) | 2.30     | 3.50 |
|                 |               |           | S. D      | 1.00  | 0.89    | (=NGAN) | 0.64     | 0.50 |
|                 |               | PSNR (dB) | 30.95     | 31.52 | (=NGAN) | 31.50   | 31.05    |      |
|                 | 0.75<br>(bpp) | I. S      | MEAN      | 2.70  | 1.70    | 1.70    | 2.30     | 2.30 |
|                 |               |           | S. D      | 0.78  | 0.46    | 0.46    | 0.78     | 0.64 |
|                 |               | PSNR (dB) | 32.10     | 32.74 | 32.73   | 32.79   | 32.39    |      |
| BABOON<br>IMAGE | 0.25<br>(bpp) | I. S      | MEAN      | 2.60  | 2.30    | (=NGAN) | 3.40     | 3.20 |
|                 |               |           | S. D      | 0.92  | 0.90    | (=NGAN) | 1.02     | 1.08 |
|                 |               | PSNR (dB) | 21.96     | 21.87 | (=NGAN) | 21.89   | 21.83    |      |
|                 | 0.5<br>(bpp)  | I. S      | MEAN      | 3.10  | 1.80    | 2.10    | 2.30     | 3.00 |
|                 |               |           | S. D      | 0.83  | 0.60    | 0.54    | 0.46     | 0.89 |
|                 |               | PSNR (dB) | 22.99     | 22.91 | 22.95   | 23.00   | 22.89    |      |
|                 | 0.75<br>(bpp) | I. S      | MEAN      | 3.40  | 1.80    | 1.70    | 2.00     | 3.00 |
|                 |               |           | S. D      | 0.49  | 0.60    | 0.64    | 0.63     | 0.45 |
|                 |               | PSNR (dB) | 24.03     | 23.63 | 23.68   | 23.87   | 24.28    |      |

표 4. 16 sub-band 분할 형태에 의한 재생 영상의 평가 결과

Table 4. Performance comparison of 16 sub-band splitting scheme.

| CONDITION    |            |           | NO WEIGHT | NGAN      | RAO       | NILL      | SAKRISON  |           |
|--------------|------------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| LEAN IMAGE   | 0.25 (bpp) | I. S.     | MEAN      | 4.50      | 2.20      | (=NGAN)   | (=NO WET) | (=NO WET) |
|              |            |           | S. D      | 1.12      | 0.60      | (=NGAN)   | (=NO WET) | (=NO WET) |
|              |            | PSNR (dB) | 28.02     | 29.33     | (=NGAN)   | (=NO WET) | (=NO WET) |           |
|              | 0.5 (bpp)  | I. S.     | MEAN      | 2.90      | 1.50      | 1.50      | (=RAO)    | (=NO WET) |
|              |            |           | S. D      | 1.04      | 0.50      | 0.50      | (=RAO)    | (=NO WET) |
|              |            | PSNR (dB) | 31.19     | 31.91     | 32.24     | (=RAO)    | (=NO WET) |           |
|              | 0.75 (bpp) | I. S.     | MEAN      | 1.40      | 1.20      | (=NGAN)   | 1.40      | 1.40      |
|              |            |           | S. D      | 0.49      | 0.40      | (=NGAN)   | 0.49      | 0.49      |
|              |            | PSNR (dB) | 33.38     | 33.90     | (=NGAN)   | 33.99     | 33.95     |           |
| BABOON IMAGE | 0.25 (bpp) | I. S.     | MEAN      | 2.40      | (=NO WET) | (=NO WET) | (=NO WET) | (=NO WET) |
|              |            |           | S. D      | 0.92      | (=NO WET) | (=NO WET) | (=NO WET) | (=NO WET) |
|              |            | PSNR (dB) | 22.11     | (=NO WET) | (=NO WET) | (=NO WET) | (=NO WET) |           |
|              | 0.5 (bpp)  | I. S.     | MEAN      | 4.00      | 2.10      | (=NGAN)   | 2.20      | 3.10      |
|              |            |           | S. D      | 1.41      | 0.54      | (=NGAN)   | 0.60      | 0.70      |
|              |            | PSNR (dB) | 22.62     | 23.15     | (=NGAN)   | 23.27     | 23.19     |           |
|              | 0.75 (bpp) | I. S.     | MEAN      | 3.20      | 1.70      | 1.30      | 1.90      | 2.10      |
|              |            |           | S. D      | 0.87      | 0.90      | 0.46      | 0.70      | 0.54      |
|              |            | PSNR (dB) | 24.43     | 23.98     | 24.03     | 24.25     | 24.78     |           |

최종적인 레벨 할당표가 같아지는 경우가 발생하였다. 이러한 현상은 저주파 영역에서의 MTF 차이가 미세하기 때문에 나타나는 현상이다. 각 결과는 부호화된 영상들에 대한 impairment scale 결과들의 평균과 분산 및 PSNR을 나타내고 있다. 표 3과 표 4의 주관적인 평가 결과들에서 알 수 있듯이, 4 가지 MTF들 중에서 Rao의 MTF가 가장 우수하게 나타났고 다음으로 Ngan, Nill, Sakrison 그리고 MTF를 적용하지 않은 경우의 순으로 평가되었다. 그림 4는 실험에서 사용한 원영상과 0.25, 0.5, 0.75 bpp 에서 인간 시각 특성을 고려한 영상과 고려하지 않은 영상을 비교하여 나타낸다. 영상의 저주파 영역을 충실히 재생하는 Rao와 Ngan의 MTF를 적용한 결과가 다른 경우들에 비하여 주관적인 기준에서 우수하게 나타났다. 영상의 관찰 각도를 6°로 한 경우에 대한 여러가지의 비교 결과에서 Rao의 MTF를 적용한 결과가 Ngan의 MTF에 의한 결과들에 비하여 blurring 현상이 약간 완화되는 것을 알 수 있었다. 대부분의 인간 시각에 대한 연구 결과에서 공간 주파수가 3-4.5 cycles/degree 에서 최대의 감도 특성을 나타낸다.<sup>12,9)</sup> 비교적 관찰 조건이 비슷한 [9]의 경우 (자연적인 동공크기, 1.42m의 관찰 거리)에 있어서 Rao의 MTF가 시각 특성의 결과와 가장 근사한 것으로

나타났다. 따라서, 본 실험의 결과는 시각 특성에 관한 연구 결과들과 상당히 일치하는 것을 알 수 있다. 한편, 표 3과 표 4의 PSNR 결과를 살펴보면 MTF 가중치를 고려하지 않은 결과 영상의 PSNR이 낮게 나오는 경우가 있다. 이러한 결과는 식(13과 식(14)의 비트 할당 방법이 준최적이기 때문인 것으로 생각된다. 따라서, [23]에서와 같은 최적 비트 할당 방법에 대한 연구가 보완되어야 할 것으로 생각된다. 저주파 대역을 세분함으로써 나타나는 부호화 결과의 영향에 대하여 살펴보면, 표 3과 표 4에서 알 수 있듯이 전체적으로 16 sub-band 분할에 의한 결과가 28 sub-band 분할에 의한 결과보다 우수하게 나타났다. 그러나, sub-band 분할 형태에 따른 영향을 자세히 알아보기 위해 0.25 bpp의 전송률에서 16 sub-band 분할 형태와 28 sub-band 분할 형태에 의한 부호화 결과를 각 전송률(bit rate)에서 객관적 및 주관적으로 우열 비교한 결과에서는 두가지 경우에 의한 결과가 비등하게 나타났다. 결과적으로 낮은 전송률에서, 신호의 저주파 대역을 세분함으로써 나타나는 비트 할당 과정에서의 잊점과 인간 시각 시스템 모델에 근접하는 잊점이 주파수 대역 분할 필터링에 의한 왜곡의 영향을 어느정도 보완한 것으로 생각된다.

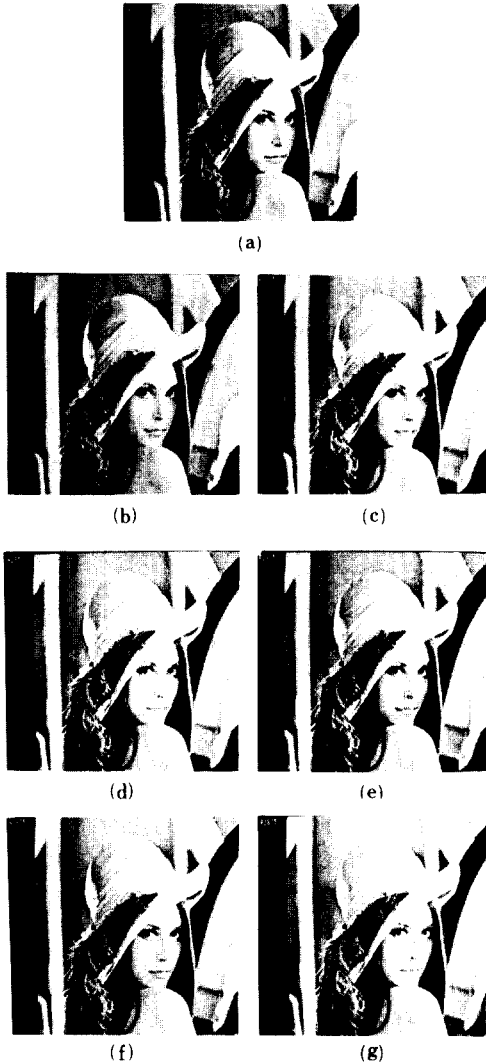


그림 4. LENA 영상의 28 sub-band 분할 형태에서의 결과 비교

- (a) LENA 원영상 (512×512)
- (b) 가중치 없음, 0.25 bpp
- (c) Rao의 MTF, 0.25 bpp
- (d) 가중치 없음, 0.5 bpp
- (e) Rao의 MTF, 0.5 bpp
- (f) 가중치 없음, 0.75 bpp
- (g) Rao의 MTF, 0.75 bpp

Fig. 4. A comparison of reconstructed images employing 28 sub-band splitting scheme.

- (a) LENA original.
- (b) No weight, 0.25 bpp.
- (c) Rao's MTF, 0.25 bpp.
- (d) No weight, 0.5 bpp.
- (e) Rao's MTF, 0.5 bpp.
- (f) No weight, 0.75 bpp.
- (g) Rao's MTF, 0.75 bpp.

#### IV. 결 론

본 논문에서는 인간 시각 시스템(human visual system:HVS)의 공간 주파수 감도 특성으로 이미 제안된 4가지의 MTF들을 sub-band 부호화에 적용하고 그 결과들을 주관적인 평가 기준에서 비교하였다.

Sub-band 분할 방식에서, 16개의 동일한 공간 주파수 대역폭을 갖는 sub-band 분할 방식에 대하여 28개의 sub-band 분할 방식에 의한 영향을 알아보았다. 각 sub-band에 대한 부호화 과정에서는 MTF로부터 각 sub-band의 가중치를 구하여 가중 평균 자승 오차를 최소화하도록 비트를 할당하였다. 시뮬레이션 결과에 대한 주관적인 평가에서, Rao의 MTF 적용 결과가 가장 우수하게 나타났고 다음으로 Ngan, Nill, Sakrison 그리고 MTF를 적용하지 않은 경우의 순으로 평가되었다. Sub-band 분할에 따른 비교에서 전체적으로 sub-band 분할 필터링에 의한 왜곡에 의해서 28 sub-band 분할 방식이 16 sub-band 분할 방식에 비하여 좋지 않은 결과를 나타냈다. 그러나 0.25 bpp의 전송율에서는 두가지 분할방식이 비등한 결과를 나타냈다. 따라서 입력 영상 신호의 낮은 주파수 대역을 세분하는 필터링에 의한 손상을 무시한다면, 주관적 및 PSNR 측면에서 성능향상이 이루어진 것을 알 수 있었으며 이에 대한 보다 심층적인 연구가 뒤따라야 할 것으로 생각된다.

#### 참 고 문 헌

- [1] M. Kunt, A. Ikonomopoulos, and M. Kocher, "Second-generation image-coding techniques," *Proc. IEEE*, vol. 73, no. 4, pp. 549-574, Apr. 1985.
- [2] F.W. Campbell and J.G. Robson, "Application of fourier analysis to the visibility of gratings," *J. Physiol.* vol. 197, pp. 551-556, 1968.
- [3] J.L. Mannos and D.J. Sakrison, "The effects of a visual fidelity criterion on the encoding of images," *IEEE Trans. Inform. Theory*, vol. IT-20, no. 4, pp. 525-536, July 1974.
- [4] N.B. Nilf, "A visual model weighted cosine transform for image compression and quality assessment," *IEEE Trans. Commun.*, vol. COM-33, no. 6, pp. 551-557, June 1985.
- [5] K.N. Ngan, K.S. Leong and H. Singh, "Cosine transform coding incorporating human visual system model," *SPIE Fiber LASE 86*, Cambridge, MA., vol. 707, pp. 165-171, Sep. 14-20, 1986.



- [6] B. Chitprasert and K.R. Rao, "Human visual weighted progressive image transmission," *ICCS '88, Intl. Conf. on Commun. Systems*, Singapore, paper no. 171.1, Nov. 1988 .
- [7] G. Karlsson and M. Vetterli, "Extension of finite length signals for sub-band coding," *Signal Processing*, vol. 17, no. 2, pp. 161-168, June 1989.
- [8] M.G. Perkins and T. Lookabaugh, "A psychophysically justified bit allocation algorithm for subband image coding systems," *Proc. ICASSP*, Glasgow, Scotland, vol. 3, pp. 1815-1818, Apr. 1989.
- [9] J.J. DePalma and E.M. Lowry, "Sine-wave response of the visual system. II. sine-wave and square-wave contrast sensitivity," *J. Opt. Soc. Am.*, vol. 52, no. 3, pp. 328-335, Mar. 1962.
- [10] D.J. Sakrison, "On the role of the observer and a distortion measure in image transmission," *IEEE Trans. Commun.*, vol. COM-25, no. 11, pp. 1251-1267, Nov. 1977.
- [11] R.E. Crochiere, S.A. Webber, and J.L. Flanagan, "Digital coding of speech in subbands," *Bell System Technical Journal*, vol. 55, no. 8, pp. 1069-1085, Oct. 1976.
- [12] M. Vetterli, "Multi-dimensional sub-band coding: Some theory and algorithms," *Signal Processing*, vol. 6, no. 2, pp. 97-112, Feb. 1984.
- [13] J.W. Woods and S.D. O'Neil, "Sub-band coding of images," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-34, no. 5, pp. 1278-1288, Oct. 1986.
- [14] H. Gharavi and A. Tabatabai, "Sub-band coding of monochrome and color images," *IEEE Trans. Circuit and Systems*, vol. 35, no. 2, pp. 207-214, Feb. 1988.
- [15] P.H. Westerink, J. Biemond, and D.E. Boekee, "Evaluation of image subband coding schemes," *Signal Processing IV*, vol. 3, pp.1149-1152, 1988.
- [16] S. Nanda and W.A. Pearlman, "Tree coding of image sub-bands," *Proc. ICASSP*, vol. 2, New York, New York, pp. 769-772, Apr. 1988.
- [17] N.S. Jayant and P. Noll, *Digital coding of Waveforms*, Englewood Cliffs, NJ: Prentice-Hall, 1984.
- [18] W.K. Pratt, *Digital Image Processing*, New York, New York: John Wiley and Sons, 1978.
- [19] A.K. Jain, *Fundamentals of Digital Image Processing*, Englewood Cliffs, NJ: Prentice-Hall, 1989.
- [20] J.D. Johnston, "A filter family desined for use in quadrature mirror filter banks," *Proc. ICASSP*, Denver, pp. 291-294, Apr. 1980.
- [21] J.O. Limb, "Distortion criteria of the human viewer," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-9, pp. 778-793, Dec. 1979.
- [22] M.B. Saches, J. Nachemias, and J.G. Robson, "Spatial-frequency channels in human vision," *J. Opt. Soc. Am.*, vol. 61, pp. 1176-1186, Sep. 1971.
- [23] P.H. Westerink, J. Biemond, and D.E. Boekee, "An optimal bit allocation algorithm for sub-band coding," *Proc. ICASSP*, New York, New York, pp. 757-760, 1988.

---

 著 者 紹 介
 

---



金 容 寬 (正會員)

1965年 3月 9日生. 1988年 2月 서울대학교 제어계측공학과 졸업 (BS). 1990年 2月 서울대학교 제어계측공학과 졸업 (MS). 1990年 3月 - 현재 서울대학교 제어계측공학과 박사과정 재학중. 주관심

분야는 디지털 신호처리 등임.

朴 涉 亨 (正會員) 第25卷 第7號 參照

현재 서울대학교 대학원 제어계측공학과 박사과정 재학중

李 商 郁 (正會員) 第25卷 第7號 參照

현재 서울대학교 제어계측공학과 부교수